

Esame SDE 20/05/2026

Note: FSDA version 8.7.10.23 or later must be installed

% Il codice chiaramente riconducibile a una generazione automatica tramite
% strumenti di intelligenza artificiale (ad esempio soluzioni prodotte
% integralmente da modelli linguistici) comporterà una penalizzazione nella
% valutazione dell'elaborato. Gli studenti sono tenuti a dimostrare una
% comprensione personale dei metodi utilizzati e a produrre codice che
% rifletta il proprio ragionamento, la propria struttura e il proprio stile
% di commento.

%%%%%%%%%%
%% DURATA: 75 minuti

Il file Matlab (script), formato m oppure mlx,
va salvato con il vostro nome e cognome (senza spazi e accenti)
e va caricato nella pagina che viene comunicata.

La votazione finale terrà conto della qualità del codice e della
sua chiara presentazione.

Indicare il nickname GitHub

Indicare (se presente) la partecipazione al seminario

Indicare (se presente) la segnalazione di refusi nel libro di testo

%%%%%%%%%%

Caricare i dati presenti nel file **univer.xlsx** in memoria come tabella utilizzando le prime due lettere del proprio cognome senza accenti

A sample survey was conducted on students' evaluation of Orientation services. Participants were asked to express their ratings on a 7-point scale:

- 1 = extremely unsatisfied
- 2 = very unsatisfied
- 3 = unsatisfied
- 4 = indifferent
- 5 = satisfied
- 6 = very satisfied
- 7 = extremely satisfied

For example

- **willingn**: level of satisfaction with the willingness of the staff
- **competete**: judgment about the competence of the staff
- **xGlobal**: global satisfaction

Creare la tabella pivot come mostrato nell'immagine seguente e denominarla **TVP**. Le variabili sulle righe sono **willingn** e **competete**, mentre la variabile sulle colonne è **xGlobal**. (10 punti)

willingn	disc_competete	[1, 5)	[5, 7]
1	[1, 4)	35	2
1	[4, 7]	4	1
2	[1, 4)	33	2
2	[4, 7]	3	2
3	[1, 4)	24	1
3	[4, 7]	8	13
4	[1, 4)	22	7
4	[4, 7]	34	30
5	[1, 4)	23	9
5	[4, 7]	44	264
6	[1, 4)	10	12
6	[4, 7]	31	571
7	[1, 4)	9	4
7	[4, 7]	9	972

```
RI=readtable('univer.xlsx');
```

Warning: Column headers from the file were modified to make them valid MATLAB identifiers before creating variable names for the table. The original column headers are saved in the VariableDescriptions property.
Set 'VariableNamingRule' to 'preserve' to use the original column headers as table variable names.

```
TVP=pivot(RI,Rows=["willingn" "competete"], ...  
    RowsBinMethod={"none" [1 4 7]}, ...  
    Columns="xGlobal",ColumnsBinMethod=[1 5 7]);
```

Creare la tabella di contingenza tra le variabili **willingn** e **competete**. (5 punti)

```
Ntable=pivot(RI,"Rows","willingn","Columns","competete","RowLabelPlacement","rownames");
```

Creare un grafico che mostri il contributo di ciascuna cella all'indice Chi-quadrato e il relativo diagramma di Pareto. (5 punti)

```
out=corrNominal(Ntable,"plots",true);
```

```
Chi2 index  
2.6189e+03
```

pvalue Chi2 index
0

Phi index
1.0963

Cramer's V
0.4476

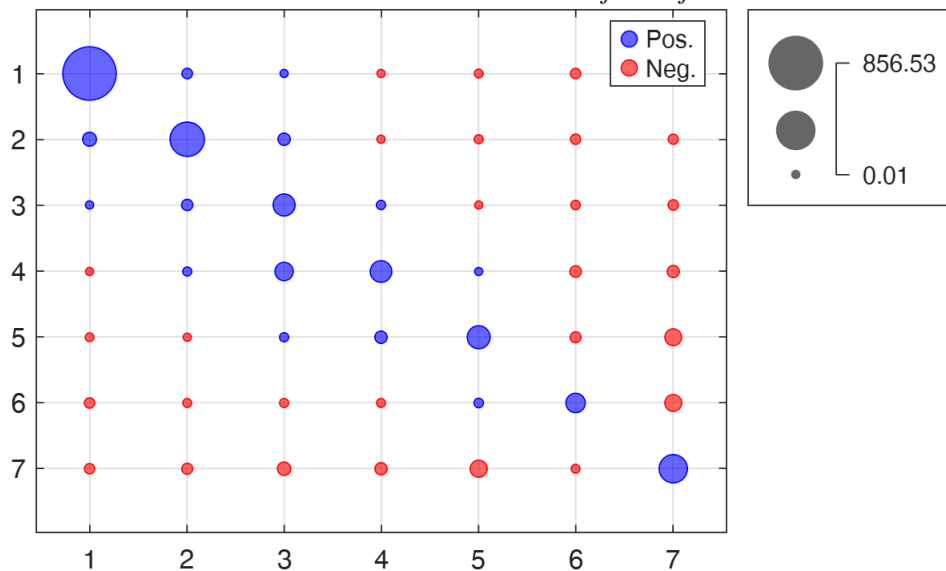
Test of H₀: independence between rows and columns

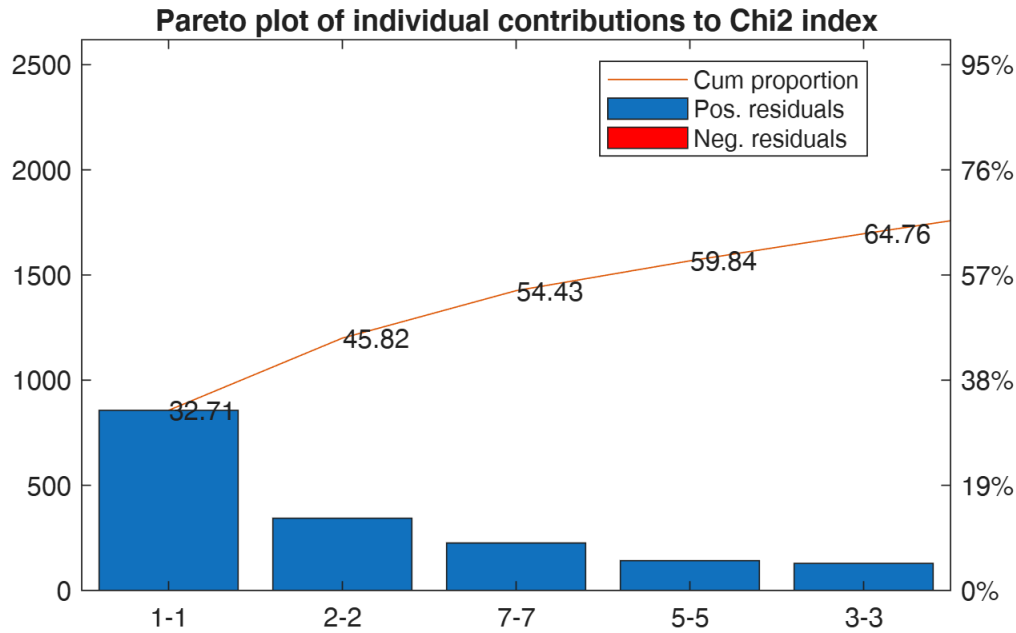
	<u>Coeff</u>	<u>se</u>	<u>zscore</u>	<u>pval</u>
CramerV	0.44756	0.010308	43.42	0
GKLambda_{yx}	0.32603	0.015025	21.699	0
tau_{yx}	0.18021	0.010068	17.9	0
H_{yx}	0.22333	0.010636	20.997	0

Indexes and 95% confidence limits

	<u>Value</u>	<u>StandardError</u>	<u>ConflimL</u>	<u>ConflimU</u>
CramerV	0.44756	0.010308	0.42736	0.46176
GKLambda_{yx}	0.32603	0.015025	0.29658	0.35548
tau_{yx}	0.18021	0.010068	0.16047	0.19994
H_{yx}	0.22333	0.010636	0.20248	0.24418

Pearson residuals² : $(\pm)(n_{ij} - n_{ij}^*)^2/n_{ij}^*$

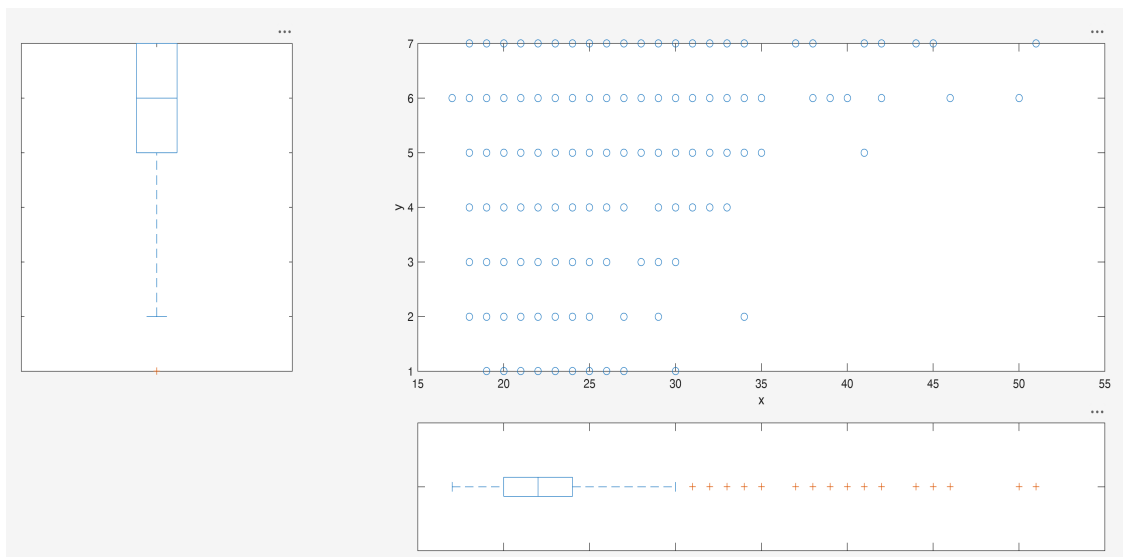




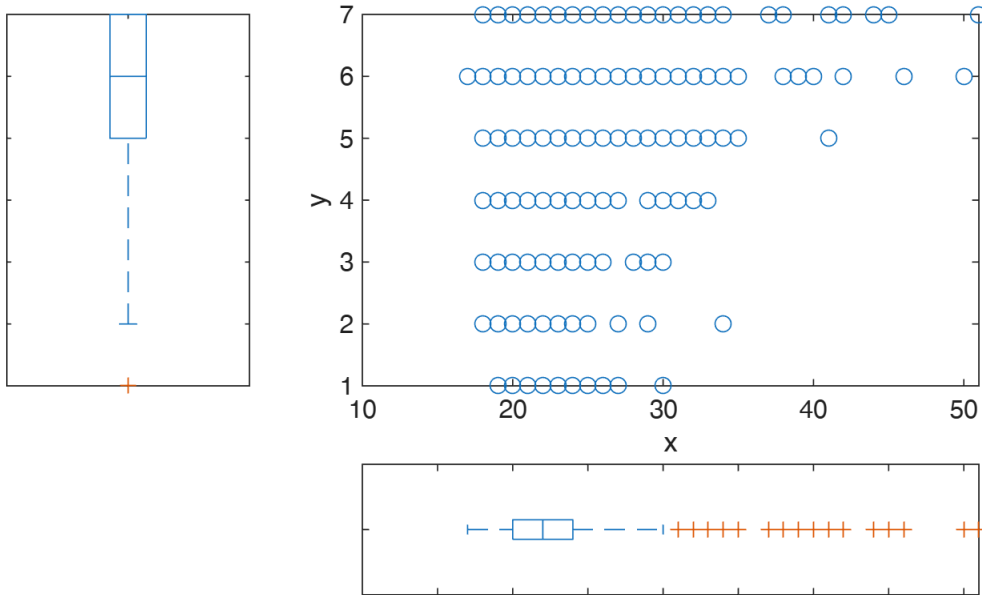
Commentare entrambi i grafici. (5 punti)

La relazione è altamente significativa. Le celle che contribuiscono maggiormente sono 1,1; 2,2 e 7,7, le quali mostrano una frequenza molto superiore a quella attesa. Queste tre celle (come mostrato dal grafico di Pareto) forniscono un contributo all'indice complessivo superiore al 50%. Le prime cinque celle, considerate congiuntamente, spiegano circa il 65% dell'associazione totale, evidenziando che i legami più forti si verificano agli estremi della scala di soddisfazione.

Calcolare il grafico riportato sotto tra le variabili **age** e **xGlobal**. Commentare gli outlier univariati. (5 punti)



```
scatterboxplot(RI.age,RI.xGlobal);
```



Per quanto riguarda l'età, gli outlier si trovano nella coda destra della distribuzione: si tratta di tutti gli individui con età superiore ai 30 anni. Per quanto concerne la soddisfazione globale, gli outlier corrispondono a coloro che hanno risposto 1 = estremamente insoddisfatto.

Il boxplot marginale dell'età (in basso) mostra diversi outlier univariati superiori, rappresentati dai simboli rossi "+". Mentre la maggior parte degli studenti ha un'età compresa tra 20 e 28 anni, esiste un gruppo distinto di studenti maturi (di età compresa tra 35 e oltre 50 anni) che vengono identificati statisticamente come outlier rispetto a questa caratteristica demografica.

Il boxplot marginale della soddisfazione globale (a sinistra) mostra che il livello di soddisfazione è generalmente elevato (la mediana si trova al livello 6). Tuttavia, individua come outlier univariati inferiori coloro che hanno risposto 1. Ciò rappresenta un caso raro di studenti "estremamente insoddisfatti", che si discosta in modo sostanziale dalla tendenza complessivamente positiva del campione.

In totale 5 righe di codice.